DOCUMENT RESUME

ED 341 103 CS 507 710

AUTHOR Dudczak, Craig; Day, Donald
TITLE Debate Philosophy Statements as Predictors of Critic Attitudes: A Summary and Direction of Research.
PUB DATE Nov 91
NOTE 16p.; Paper presented at the Annual Meeting of the Speech Communication Association (77th, Atlanta, GA, October 31-November 3, 1991).
PUB TYPE Information Analyses (070) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Comparative Analysis; *Content Analysis; *Debate; Higher Education; Interrater Reliability; Judges; *Persuasive Discourse; *Public Speaking; Questionnaires; Research Methodology; Research Needs
IDENTIFIERS Cross Examination Debate Association; *Debate Theory; *Debate Tournaments; National Debate Tournament

ABSTRACT

Philosophy statements have been used in the National Debate Tournament (NDT) since the mid-1970s and the Cross Examination Debate Association (CEDA) National Tournament since its 1986 inception. The statements should help debaters adapt to critics' expressed preferences. Moreover, philosophy statements can guide the study of argumentation theory and practice. Philosophy statements have been examined through: (1) self-report instruments completed by debate critics; (2) content analysis of judge philosophy statements; and (3) examination of CEDA and NDT debate critics' ballots. A fourth, "integrated," approach combines two or more sources of data and/or methods of data analysis. A study using content analysis in comparing NDT judge philosophy statements with ballots found high consistency between the two. Lower levels of consistency were found in four studies which compared the debate critics' professed preferences with their expressed ballot behavior and used survey instruments in combination with content analysis to evaluate debate critic behavior. Further research should examine the suggestion that judge philosophy statements have substantially higher predictive power than do survey questionnaires. The willingness of critics to employ paradigms other than their expressed preferences also bears study. It is also necessary to determine whether researchers' measurement instruments are reliable. (One figure is included; 21 references are attached.) (SG)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED341103

# Debate Philosophy Statements as Predictors of Critic Attitudes: A Summary and Direction of Research

Craig Dudczak and Donald Day

Syracuse University

Speech Communication Association Conference

Atlanta, Georgia

November 1991

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Craig Dudczak

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS507710

Running Head: Debate Philosophy Statements in CEDA



BEST COPY AVAILABLE

Debate Philosophy Statements as Predictors of Critic Attitudes:
A Summary and Research Direction

Philosophy statements have been used in academic debate since the mid-1970's at the National Debate Tournament (NDT), and at the Cross Examination Debate Association (CEDA) National Tournament since its inception in 1986. Philosophy statements assume that debate critics better formulate their decision criteria by articulating them. Once articulated, these statements should better enable debaters to adapt to their critics' expressed preferences. Moreover, treating debate as a laboratory for of applied argumentation, philosophy statements serve as a guide for the study of argumentation theory and practice.

While the use of philosophy statements has been generally accepted in the CEDA debate community,[1] little evidence exists confirming that these assumptions are true. If philosophy statements do not reflect the decision criteria actually applied by critics in debate rounds, their utility may be called into question.

## REVIEW OF LITERATURE

A number of studies have evaluated critics' paradigm preferences in NDT through the use of self-report instruments (Cox 1974; Cross & Matlon 1978; Thomas 1977) and in CEDA (Buckley 1983; Lee, Lee & Seeger 1983). These reports typically asked subjects to indicate their preferred decision paradigm and respond to situations which might occur in a debate. These early

studies did not attempt to establish whether preferences expressed by debate critics were actually used by them in judging debates. Because of their exclusive reliance upon self-report instruments, these attitude surveys are subject to an array of reporting biases. As Carlsmith et al (1976) caution, subjects may be "either unable or unwilling to comment on ongoing processes." (71) Without behavioral confirmation of reported preferences, the accuracy of such instruments remains an open question. Nevertheless, these reports have value in identifying variables for subsequent research.

A second approach to identifying decision preferences has been to analyze judge philosophy statements through some form of content analysis. Brey (1989; 1990), for instance, analyzed the content of CEDA philosophy statements to summarize accepted and disliked tactics and arguments in CEDA debate.

Although philosophy statements are a type of self-report, they differ from survey instruments. Survey instruments typically pre-structure respondents' answers to conform with options offered by the researcher. In other words, respondents' choices are dictated by the instrument. Content analysis (of philosophy statements), however, begins with a view of reality held by the subject and attempts to conform that world to the analytic scheme of the researcher (Holsti 1969; Krippendorff 1980; Weber 1985).

The attempt to conform the respondents' view of reality to the research scheme presents its own set of limitations. The

general limitations surround how one interprets the meaning present in the written artifact (philosophy statement) provided by the subjects. While the subjects provide a more or less unstructured response of their preferences, the analytic scheme super-imposes the researcher's expectations as a filter upon these responses. In other words, the researcher is likely to find what s/he is looking for. Delia and Grossberg (1977) remind us that the interpretation of meaning should reflect the subjects' reality.

The specific limitation of using debate philosophy statements is that they are also largely structured by a set of questions critics are asked to answer. The issues critics are asked to address, often with a space limitation for their response, tends to determine the content of their statements. [2] Hence, general interpretive biases created by the researcher's expectations are likely to be accentuated by pre-structuring much of the subject's responses.

Nevertheless, even with these limitations, content analysis of philosophy statements yields valuable information. Philosophy statements tend to be more generalized, enduring constructions of a critic's perspective. Rather than generating situation-specific responses, the philosophy projects a dispositional attitude of preference.

Researchers have not limited themselves to analysis of self-reports of professed preferences. A third approach employed in the study of judging behavior has been to look at the artifacts

generated by debate critics. Hollihan, Riley, and Austin (1983), for example, used content analysis of NDT and CEDA ballots to determine thematic "visions" embraced respectively within these two debate formats. Analysis of behavior (as reflected through ballots) avoids the reporting biases associated with survey responses.

There are also limitations associated with only analyzing the artifacts provided. The interpretive limitations attendant to content analysis remain, for instance. The researchers still super-impose their construction of reality upon the artifact to make it meaningful. Moreover, this becomes more likely since the subject's intent is never solicited. Limiting one's inquiry to only the behavioral artifact without knowledge of the critics' prior attitudes makes the researcher's interpretive frame paramount. Subjects are not asked what they had in mind when they wrote their comments: The researcher presumes to know best.

An additional limitation becomes likely. The ballot, as an artifact of behavior, may reflect a dispositional preference of the critic, a response to the situation created by the particular debate, or some combination of both. One cannot know whether ballot comments reflected critic preference or circumstances unique to debate rounds. Analysis of multiple ballots from the same critic is required to minimize the alternative explanations for a critic's responses.

A final approach to the study of judge behavior may be called the "integrated" approach. Such research attempts to

combine two more sources of data and/or methods of data analysis. Comparing preferences expressed through a philosophy statement with actual behavior compensates for some of the limitations attendant to viewing each separately. Similarly, the use of survey research in concert with content analysis can yield complementary findings which are more valid than those obtained using either alone (Paisley 1969; Webb and Roberts 1969).

There were five research reports we consider to be "integrated" in their approach. The first, by Henderson and Boman (1983) compared judge philosophy statements with ballot artifacts using content analysis to analyze each. They reported high consistency (83.5%) between a set of NDT judge philosophy statements and corresponding ballot comments. Dudczak & Day (1990; Day and Dudczak 1991) have previously questioned their analytic procedures. The use of a single ballot for most critics analyzed makes the representativeness of the ballot artifacts suspect.[3]

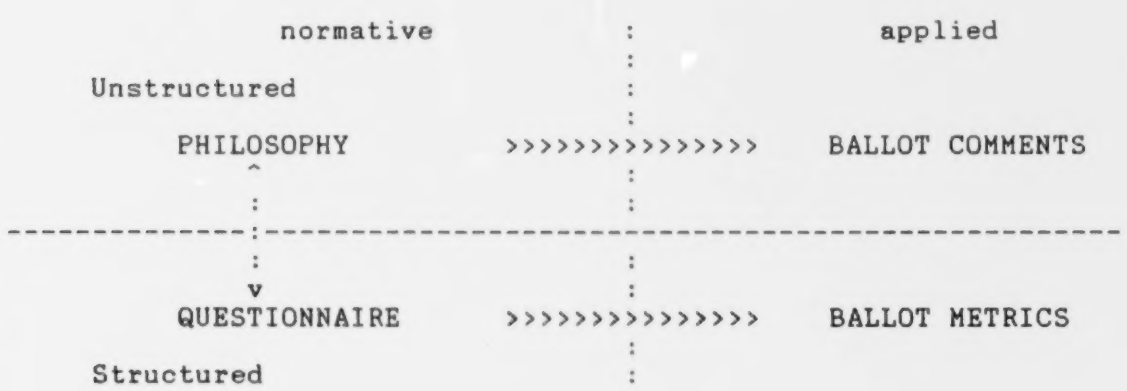
As an integrated study, however, Henderson and Boman make an important contribution. Theirs was the first study to compare the professed preference of debate critics with their subsequent behavior (as expressed through ballots).

Four studies reported by Dudczak and Day (1989a; 1989b; 1990; Day & Dudczak 1991) have compared both the preferences professed by debate critics with their expressed ballot behavior as well as used survey instruments in combination with content analysis to evaluate debate critic behavior. The integrated

design for this research program is represented in Figure 1. The one instrument and two work products used in the study may be visualized in a two-by-two table. Both the philosophy and questionnaire are normative -- "ought" -- documents; the ballots are applied documents. The philosophy and comment portions of ballots are unstructured; the questionnaire and template (top) portions of ballots are structured.

FIGURE 1

Construct and technique matrix of tools in the study




Judges' preferences were determined through two independent measures; philosophy statements and survey questionnaires. The use of multiple measures allows the assessment of predictive validity. The use of measurement instruments over a series of applications allows reliability calculation. The results of three experiments using a non-regional sample plus the results of a regional pilot study are reported here.

Dudczak and Day (1989a) found lower consistency (54.8%) among debate critics than Henderson and Boman (1983) reported. Dudczak and Day (1989a) also reported that several clusters of paradigms were correlated with decision criteria cited in critics' ballots. A secondary analysis of Dudczak and Day's pilot data (1989b) sought to isolate differences among traditional paradigms. Paradigm boundaries were found to be porous and unreliable.

Unlike the earlier work by Dudczak and Day (which included only data from the Northeast), their 1990 non-regional study included tournaments from across the U.S. Their first two experiments replicated the previous pilot effort, investigating three research questions and nine hypotheses. Results showed little reliability for questionnaires as predictors of critics' ballot behavior. The 1990 experiments by Dudczak and Day showed limited association between professed paradigms and subsequent ballot behavior, and indicated that traditional paradigms largely overlap each other. In fact, the non-regional study indicated less consistency between professed beliefs and actual ballot behavior than had been observed with purely regional data.

The latest experiment by Day and Dudczak (1991) compared variables on questionnaires to corresponding variables on philosophies, to evaluate the degree to which the instruments measure similar aspects of critic preference. That experiment showed little similarity between the two instruments. It also demonstrated that inconsistencies between professed and actual

behavior noted in earlier work were not an artifact of intrasample cancellation due to data aggregation: critics were inconsistent individually, not merely as a group.

## Direction for Further Research

1) Which instrument (philosophy statement or survey questionnaire) better predicts how debate critics will behave?

Day and Dudczak (1991) sought to establish the degree to which questionnaires and philosophy statements map to each other (i.e., the extent to which they made consistent predictions). To the extent they vary considerably in their predictions (as reported in Day & Dudczak 1991), it is likely that a) one has a higher level of predictive validity than the other, (b) both are equally predictive for varying reasons, or (c) both are equally non-predictive for varying reasons. In addressing the problem of instruments' predictive validity, evidence reported by Dudczak & Day's regional pilot study (1988a) indicates that judge philosophy statements have substantially higher predictive power than do survey questionnaires. Further research should establish which of these alternatives is most probably true.

2) Are debate paradigms meaningful indicators of critics' decision making behavior?

Dudczak and Day (1990) have commented previously that paradigms are "porous and unreliable." Few distinctive elements discriminating among the several paradigm could be found when

they were correlated with ballots comments (Dudczak & Day 1989b; 1990). Further, the widespread willingness (94%) of critics to employ a paradigm other than their professed preference (Dudczak & Day 1989a) suggests that a) paradigms are not meaningful predictors of subsequent behavior, or b) the paradigms are not well understood by the critic judges who employ them.

Determining which alternative is more likely true requires an assessment of the "accuracy" by which paradigms are understood. Accuracy is that dimension of reliability by which a behavior is assessed against a standard or norm (Weber 1985). While a literature describing the characteristics of the several paradigms exists, there is no certification of critics who use them. If critics' explanations of their preferred paradigm corresponded with the standard for the paradigm (as established by its literature), then indirect support for the first explanation would be offered. However, if critics' explanations were inconsistent with their preferred paradigm, then direct support for the second explanation would be available.

3) Are the measurement instruments reliable?

While this review of literature has been critical of the instruments, design and procedures employed by several previous studies, the ongoing research project conducted by Dudczak and Day is not immune to the same criticisms. All of the limitations specific to the individual research methods still apply to

whoever uses them. The integrated approach compensates for the more severe limitations, but can never completely eliminate them.

Specific means of improving reliability focus on the following areas: a) obtaining "critical" cell size for the quantitative analysis, and b) improving inter-coder reliability for content analysis.

The researchers have been limited in the number of subjects available to the study by self-imposed constraints. One example has been the establishment of a threshold minimum of six ballots written by a critic before s/he would be included in the subject pool. This threshold was set based on the assumption that too few ballots from a critic would distort comparisons between professed preferences and actual behavior. Situational variables unique to a single round would creates anomalies between what the critic believed and the round s/he was forced to evaluate. However, the previously discussed willingness of critics to abandon their professed paradigm preference suggests the exception is actually the rule. Lowering the threshold by a single ballot would greatly increase the number of usable subjects.

The other primary element is whether the computation of inter-coder reliability would be could be improved by either revised instruments or better coding protocols. Coding forms used for Dudczak and Day's first two non-regional experiments (1990) were further expanded to include new discriminants; the coding category description form developed for those experiments

also was revised to minimize ambiguity and overlap among discriminants.

However, we're not entirely convinced that the conventions normally employed for calculating inter-coder reliability should be employed. Standard references for reliability calculation (Scott 1955) and threshold acceptability (Krippendorff 1980) are more liberal than the methods employed in the several studies. Normal calculations for reliability allow the mutual non-selection of a coding category to be considered as "agreement" between coders. We believe this artificially boosts the appearance of reliability, but fails to represent its true dimension. We will re-examine coding categories for exclusivity and exhaustiveness, but our concerns about conventions for reliability calculation will continue to direct us toward conservative estimates for reliability.

## Conclusion

Research investigating the relationship between debate critics professed beliefs and their actual behavior has been a recent phenomenon. It should be pursued to determine whether the assumed values for judging philosophies are actually confirmed. If judging philosophies do not have a strong relationship to how a debate critic employs decision criteria in adjudicating debates, the pedagogical justification for their continued use would need to be seriously reconsidered.

## References Cited

Brey, J. (1989). "A Descriptive Analysis of CEDA Judging Philosophies Part I: Definitive Acceptance or Rejection of Certain Tactics and Arguments." CEDA YEARBOOK 10: 67-77.

Brey, J. (1990). "An Analysis of CEDA Judging Philosophies - Part II: Accepting Certain Tactics and Arguments with Reservations." CEDA YEARBOOK 11: 72-79.

Bryant, M. (1983). "A Content Analysis of Selected NDT and CEDA Transcripts." Unpublished. Paper presented at the Third SCA/AFA Summer Conference on Argumentation. Alta, Utah.

Buckley, D. (1983). "A Comparison of Judging Paradigms." ARGUMENT IN TRANSITION: PROCEEDINGS OF THE THIRD SUMMER CONFERENCE ON ARGUMENTATION (Hereinafter ARGUMENT IN TRANSITION). Eds. Zarefsky, D., Sillars, M., and Rhodes, J. Annandale, VA: Speech Communication Association, 858-870.

Carlsmith, J., P. Ellsworth, and E. Aronson (1976). METHODS OF RESEARCH IN SOCIAL PSYCHOLOGY. Reading, MA: Addison-Wesley.

Cox, J. (1974). "A Study of Judging Philosophies of the Participants of the National Debate Tournament." JOURNAL OF THE AMERICAN FORENSIC ASSOCIATION (JAFA) 11: 61-71.

Cross, J., and R. Matlon (1978). "An Analysis of Judging Philosophies in Academic Debate." JAFA 15: 110-123.

Day, D., and Dudczak, C. (1991). "Debate Philosophy Statements As Predictors of Critical Attitudes". Unpublished. Paper presented to the 7th SCA/AFA Summer Conference on Argumentation. Alta, Utah.

Delia, J., and L. Grossberg (1977). "Interpretation and Evidence." WESTERN JOURNAL OF SPEECH COMMUNICATION 41: 32-42.

Dudczak, C., and D. Day (1989a). "A Profile of CEDA Debate Critics," SPHERES OF ARGUMENT: PROCEEDINGS OF THE SIXTH SCA/AFA SUMMER CONFERENCE ON ARGUMENTATION. Ed. Gronbeck, B. Annandale, VA: Speech Communication Association, 427-433.

Dudczak, C., and D. Day (1989b). "A Taxonomy of CEDA Debate Critics". Unpublished. Paper presented to the Speech Communication Association Conference, San Francisco. (ERIC # ED 318 071).

Dudczak, C., and D. Day (1990). "The Impact of Paradigm Consistency on Taxonomic Boundaries in CEDA Debate". CEDA YEARBOOK 12 (in press). Paper presented to the Speech Communication Association Conference, Chicago. (ERIC # ED 328 925).

Henderson, B., and D. Boman (1983). A Study to Determine If Debate Judges' Philosophy Statements are Consistent with Their Later Related Ballot Statements." JAFA 19: 191-198.

Holsti, O. (1969). CONTENT ANALYSIS FOR THE SOCIAL SCIENCES AND HUMANITIES. Reading, MA: Addison-Wesley.

Krippendorff, K. (1980). CONTENT ANALYSIS: AN INTRODUCTION TO ITS METHODOLOGY. Beverly Hills, CA: Sage.

Lee, R., Lee, K., and Seeger, M. (1983). "Comparison of CEDA and NDT: Differences in Program Characteristics and Forensic Director Attitudes". ARGUMENT IN TRANSITION. Eds. Zarefsky, D., Sillars, M., and Rhodes, J. Annandale, VA: Speech Communication Association, 745-857.

Paisley, W. (1969). "Introduction." THE ANALYSIS OF COMMUNICATION CONTENT. Gerbner, G. et al., eds. New York: Wiley.

Scott, W. (1955). "Reliability of Content Analysis: The Case of Nominal Scale Coding." PUBLIC OPINION QUARTERLY 19: 321-325.

Thomas, D. (1977). "Analysis of Judges and Debaters at NDT." Unpublished. Data collected from the 1977 National Debate Tournament. Used with permission.

Webb, E. and K. Roberts (1969). "Unconventional Uses of Content Analysis in Social Science." THE ANALYSIS OF COMMUNICATION CONTENT. Gerbner, G. et al, eds. New York: Wiley.

Weber, R. (1985). BASIC CONTENT ANALYSIS. (Sage University Paper No. 49) Beverly Hills, CA: Sage.

## Endnotes

1. There are exceptions. One critic responded to the 1990 CEDA Judging Philosophy request by disparaging the utility of the statement. (See Todd Graham, 1990 Judging Philosophy Booklet.)

2. One judge criticized the absurdity of the CEDA Judging Form requesting critics to answer a page full of questions about their philosophies while limiting them to a single page. (See James J. Unger, 1990 CEDA Judging Philosophy Booklet.)

3. Henderson and Boman failed to conform to several validity and reliability standards. Primary is exhaustiveness in the content analytic scheme. Only items which appeared on both the philosophy statement and ballot were coded for consistency. The non-use of a category expressed on the philosophy because it wasn't used for the ballot studied is ambiguous. Was its absence related to its inappropriateness for the round in question or the failure of the critic to apply his/her standard?